DOCUMENT RESUME

ED 292 794 SP 030 104

AUTHOR Packard, Richard D.; Dereshiwsky, Mary I.
TITLE Evaluation Research. Study of the Effects of a Career Ladder Intervention Program with Focus on the Production & Outcomes in Student Achievement.
PUB DATE 30 Apr 88
NOTE 14p.; Paper presented to the Arizona Educational Research Association for the Conference on Partnerships in Education (April 30, 1988).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Descriptive (141)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Academic Achievement; *Career Ladders; Elementary Secondary Education; *Prediction; *Program Effectiveness; *Program Evaluation; *Research Methodology; State Programs

ABSTRACT

In 1984, Arizona governmental officials began studying a social problem plaguing both the state and country for many years: "Does a teacher incentive program enhance the profession and student achievement?" Research and evaluation has been underway for over 3 years and is focusing on the parameters of a model which will be reported to policy makers in 1989 to determine future state-wide expansion. Since student achievement is a key issue of policy research, the following methodologies for its study have been devised: (1) pre-test, post-test, gain score elements; (2) multivariate regression model; (3) canonical correlation; and (4) qualitative matrix paradigm. The first goal of the analysis will be to assess the magnitude and direction of gain scores for selected student achievement measures within individual districts. Secondly, a predictive model will be developed within selected districts which will attempt to link multiple district-specific student achievement and teacher performance measures, while controlling for possible pre-existing extraneous influences on student performances. (Author/JD)

**********
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
**********

ED292794

# EVALUATION RESEARCH

## Study of the Effects of a Career Ladder Intervention Program With Focus on the Production & Outcomes in Student Achievement

---

by

Dr. Richard D. Packard, Project Manager

The Arizona Career Ladder Research & Evaluation Project

&

Dr. Mary I. Dereshiwsky, Statistical Research Specialist

Center for Excellence in Education - Northern Arizona University

Box 5774 - Flagstaff, Arizona - 86011

---

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

R. D. Packard

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Research Document Submitted to the

Arizona Educational Research Association

for the Conference on Partnerships in Education

April 30, 1988

SP030104

## ABSTRACT

In 1984, Arizona governmental officials began studying a social problem plaguing both the state and country for many years: "Does a teacher incentive program enhance the profession and student achievement?" Research and evaluation has been underway for over three years and is focusing on the parameters of a model which will be reported to policy makers in 1989 to determine future state-wide expansion. Since student achievement is a key issue of policy research, the following methodologies for its study have been devised: (1) pre-test, post-test, gain score elements, (2) multivariate regression model, (3) canonical correlation and (4) qualitative matrix paradigm.

# EVALUATION RESEARCH:

## Study of The Effects of A Career Ladder Intervention Program With Focus On The Production & Outcomes in Student Achievement

## INTRODUCTION

### Overview

In 1984 the Arizona Executive and Legislative branches of government began moving on a social problem which had been plaguing the country, as well as the state, for many years. The major issues had to do with concerns over student achievement (or lack of it), and the number and quality of teachers interested in moving into, and remaining in, the "profession." The basic problem was to determine if teachers could be rewarded based on job performance or competencies (tied to student outcomes), rather than solely on years of experience and assumed developmental effects on students.

In 1985 the Arizona Legislature passed into law a bill which would establish a social intervention program in the interest of enhancing teacher performance and improving student achievement. Therefore, the "Career Ladder Teacher Incentive Program" had been effected. Interested districts were to apply for approval to enter into the five-year pilot test program; the Joint Legislative Committee on Career Ladders was established to direct it; and "evaluation research" was to be carried out by the Center for Excellence in Education at Northern Arizona University.

The pilot project is presently in its third year of testing, and the final summative evaluation is to be presented to the Legislature in the Fall of 1989. At this time, the 15 school districts remaining in the test are involved in ongoing program evaluation and improvement. Their progress will continue to be monitored throughout the five-year period.

### Evaluation Research

The program is being researched and evaluated through a field study technique called "evaluation research." Evaluation research is commonly referred to as "Policy Research," because

it is often used to help official decision makers in meeting some particular social need or purpose. According to Baker (1988), evaluation research experienced a considerable increase in use in the 1970s, due to the need to study and evaluate several social programs which evolved in the 1960s.

Evaluation research is not a new or different research methodology; because it makes use of several common social and behavioral scientific approaches. These paradigms provide a study design or structure to assist in addressing and understanding the effects of some societal event, problem or program. It can be described as a type of design which directs an objective and empirical study, using the most appropriate social and behavioral techniques available to the field. In many cases, it helps policy makers refine definitions or determine their specific goals before they can clearly understand and develop sound policy.

Social programs (like the comprehensive career ladder intervention program) are quite often very costly, requiring considerable funding support from taxpayers. Government officials have understandably been extremely concerned about assessing program effectiveness prior to hasty implementation or continuation. Therefore, the Center for Excellence in Education was selected to conduct the five-year study and to report program results to legislative decision-makers.

## The Purpose

Policy and evaluation research is a very comprehensive undertaking. It is the responsibility of research and evaluation specialists to view all related components which impact on the focused intent of the public policy. For conceptualization and direction, one must identify and draw parameters around those elements which appear to have the greatest effect in describing and evaluating the issue. While the career ladder legislative guidelines and purposes are generally clear, the intent of this paper is to identify the specifics of what is to be evaluated and to determine if the program policy is resulting in acceptable outcomes.

Naturally, with any type of long-term evaluation research, the purpose (or the program's major objective or intent) is continually being refined and focused. For example (in general), the five year pilot test career ladder program is to determine: (1) if the business of education is positively influenced when teachers are compensated based on instructional performance, rather than years of

experience, (2) if the program enhances recruitment, retention and motivation of high-quality teachers, (3) if it will develop and improve teacher performance in the classroom and (4) most importantly, if it will, in fact, improve student achievement.

These concepts are quite easy to generate and accept as key social, economic and philosophical issues. It is extremely difficult, though, to determine the several specific program components and factors which combine to influence desirable ends (or dependent variables).

## Identification of Program Model Components

Several interacting variables dependent on the effects of the career ladder program have been identified (Packard, 1988, March 24). Of course, the overriding focus is, "What does career ladder program intervention do for development of teaching performance; and what are its effects on improvement in student achievement?"

The twelve factors which have been formulated into a relational model are presented in Appendix A of this document. In conjunction with the pilot district network, a separate study of each of the elements is being planned or has already been implemented. For example, each of the components of "Legislative Guidelines," "Support of Governing Boards," Assessment of District Readiness Levels," "Program Designs," "Essential Elements of Career Ladder Models," and "Production & Outcomes in Student Achievement," will be described, researched and evaluated.

Each one of the elements ties into the others in very important ways. They could be visualized as a "web of relationship," each exhibiting an importance of its own, but also being an essential and related part of the total universe of concepts.

## Research Procedures

The career ladder project presently includes study (and evaluation feedback) in 15 pilot test districts, involving approximately 10,000 educators and their students and 12 million dollars in incentive funds. The districts range from small to large, are located both in rural and urban settings and include a variety of ethnic backgrounds. Two of them are located on the Navajo Reservation.

Data is being collected through a variety of procedures, most of which involve validated and reliable survey techniques. A "Network Committee" (and "Task Force" within the Network), has

been developed which is composed of representation from all districts. These groups are extremely important to the ongoing process of formulation of research questions and in cooperating and making recommendations for adequate procedures of data collection. Computerized analysis of the data is currently being conducted with the "Honeywell" mainframe through the use of the SPSSx statistical package located in the Computer Center at Northern Arizona University.

Findings to date have been presented to the Legislature as well as at professional conferences in this country and Europe, and are being published through a variety of clearinghouse resources, journals and conference proceedings. The more than 20 documents which have been developed by Packard, Dereshiwsky, Bas-Isaac and others (1985; 1986; 1987 and 1988) will not be listed in the reference section, but may be secured through contacting the research project at the following address: Dr. Richard D. Packard, Director of The Arizona Career Ladder Research & Evaluation Project, Box 5774, Northern Arizona University, Flagstsff, Arizona 86011, or phone (602) 523-5852.

## Research & Publication Objective

The research project is in the process of developing a series of documents focusing on the overall data base of concepts identified as relevant to each element of the model (see Appendix A).

The following particular relational component is contained within the very important element of Production and Outcomes in Student Achievement. Student achievement is the most crucial dependent variable of the seven essential elements of the career ladder model (for a separate diagram of the seven "essential elements," please see Appendix B; Packard, 1988, April 18). The intent of this paper is to describe perceptions and projections about the student achievement module.

## Production & Outcomes in Student Achievement

### Rationale

By providing incentives for outstanding classroom performance, the goal of career ladders is to de-emphasize accumulation of college credit and years of experience as primary reward criteria. By doing so, it is hoped that superior teacher performance, as explicitly incorporated into the reward

structure, will in fact result in greater student achievement. Rigorous exploration of the extent of the relationship between well-developed, locally accepted measures of teacher performance and student achievement would constitute a major contribution to the current state of evaluation policy and research.

## Predictive Achievement Model Design and Methodology

The following sections describe four district approaches for identifying and measuring gains in student achievement. The first of these assesses the direction and magnitude of change scores across time in locally selected student achievement scores. The second is a complex predictive model, a specific adaption of Helmstadter's (1987) initial research, which will assess the direct association between individual, specific teacher performance and student achievement measures, while attempting to control for as many extraneous influences on student performance as can be identified locally by users. Thirdly, the relationships which emerge from this predictive model will be cross-validated via profile analysis. Finally, qualitative matrix modeling methods will be used to array graphically clusters of open-ended opinion responses.

Gain-Score Assessment : Pre- & Post-Measure Analysis. Districts will be asked to identify a student achievement measure, or measures, which will be administered twice during the school year. Dereshiwsky and Packard (1988) have stated, "The direction and magnitude of average student achievement (gain scores, or difference between pre- and post-test) can be assessed using a matched-pairs t-test, in the case of a single student achievement measure."

Multivariate Predictive Model: Linking Teacher Performance & Student Achievement Measures. Some districts will use multiple measures of student achievement. These could take the form of a series of tests in individual subject areas, for example. The important thing is for each district to select whatever measure, or measures, are customary or most appropriate for its own purposes. In this case, these would be a vector of change (gain) scores, instead of a single gain score. The research question, however, would be exactly the same as in the previous instance. That is, is the magnitude and direction of the set of average achievement change

scores statistically significant? The multivariate analog of the matched-pairs t-test, or Hotelling's $T^2$, would be applied.

Selected districts are attempting to develop a complex predictive model which would link multiple teacher performance measures with multiple student achievement measures. As before, each district will choose its own most relevant performance and student achievement measures.

A multivariate regression model will serve as the framework for building this model. The set of multiple student-achievement scores will constitute the simultaneous dependent variables. The predictors, or regressors, will fall into two basic categories:

a) The set of multiple teacher performance measures, as judged appropriate by a given district (e.g., locally developed administrator- and peer-evaluation surveys).

b) A set of district- and student-specific demographic variables, which are judged to affect student achievement, and yet not directly controllable by teacher performance (e.g.,ability level; ethnicity of students; average per capita income in district).

By incorporating the above "uncontrollable" social/behavioral variables and other extraneous factors, their independent effect upon student performance can be partialled out and numerically isolated. As a result, one may obtain a more precise linkage between the teacher performance measures and the student achievement measures.

The algebraic form of the predictive model is as follows:

$$y=xB+E$$

Where y: matrix of multiple student performance measures

x: matrix of multiple predictors, or independent, variables (both teacher performance and identifiable social/behavioral factors, as discussed above)

B: matrix of associated regression coefficients corresponding to the above independent variables

E: matrix of error terms, or all other extraneous factors which have not been explicitly incorporated into the model.

To see whether there is a significant relationship overall between the two sets of variables (student achievement and predictors), Wilks Lambda ($\Lambda$) will be calculated. This is a commonly used multivariate measure which is an inverse function of F.

Next, the overall regression will be disaggregated, in order to identify which individual student achievement variables (both teacher performance and social/behavioral) are strongly associated. The multivariate regression will, therefore, be followed by a series of univariate regressions: testing individual student performance measures separately.

For each teacher performance measure, a partial correlation coefficient will be computed. This shows the correlation between that predictor and the student achievement measure, regardless of the predictor's possible intercorrelation with the other predictors specified in the regression model. This statistic is critical in the case of the teacher performance variables. It allows removal of the effects of social/behavioral variables separately, thereby isolating the incremental correlation of student performance measureswith individual teacher achievement variables. In addition, the customary outputs of univariate regression models will be reported, such as F-tests, multiple $R^2$s and adjusted $R^2$s, and t-tests of significance of the individual partial regression coefficients. In like manner, partial regression coefficients will be computed and their magnitude evaluated by means of t-tests.

The original multivariate model can be adapted to show any change in factors across successive years; in effect, it can become an econometric or time-series model. In this manner, groups of students can be tracked across time and changes in their achievement isolated.

Canonical correlation analysis. The above multivariate regression results will next be cross-validated by a canonical correlation analysis. The goal is to generate a "profile analysis," or vectors of loadings which will indicate which student achievement measures are strongly associated with which teacher performance and demographic variables. The loadings, like the partial correlation and regression coefficients, indicate the individual effects of the teacher performance and demographic factors, independent of their underlying covariability.

Supplementary covariance analysis. A second, more mathematically sophisticated way to help control for extraneous factors (e.g., teachers' necessity to take assigned classes "as is," as

opposed to being able to assign students to classes randomly) is by covariance analysis. Where feasible, student achievement pre-test measures may be used as covariates, along with any other pertinent and readily available data such as I.Q. scores. This would provide an "after-the-fact" correction for lack of random subject assignment, by equalizing group means for pre-existing conditions.

**Qualitative data analysis**. Certain student achievement measures are less amenable to precise operationalization; for example, art and music skills acquisition. For these behaviors, a teacher might make evaluative comments on the direction and the degree of mastery for each student.

Qualitative matrix modeling would allow grouping of these evaluative comments, as well as the changes in student behaviors which have been observed, for such skills. Comments can be clustered to reflect the most frequently occurring dimensions of change in student performance.

**Analysis of ranked data**. For other student behaviors, a ranking scale might be appropriate for teacher ratings of student performance. Nonparametric equivalents of the matched-pairs t-test and the correlation coefficient would be used to assess direction, change and association in measurements when the data is in the form of ranks rather than absolute measures.

**Summary comments on data analysis & methodology**. There is a clear advantage to conducting the above research on a district-by-district basis: relevance. Each district will be free to choose its own familiar, customary instruments for measuring teacher performance and student achievement. This is far preferable to pre-imposing a researcher-selected but completely unfamiliar instrument upon participating districts. A crucial part of the ultimate success of any research project is its perception by both subjects and users. Giving individual districts a major say in selecting these measures helps insure that the model which is developed will be readily understood, as well as actually used in their decision-making. For applied research cannot exist in a vacuum; it is designed and executed for a practical purpose, to answer a question and/or fulfill a specific need.

A second reason for a district-specific model is that the extraneous (demographic) factors will, of necessity, be unique to each particular district. Therefore, each district is the most reliable source of ideas as to what these relevant extraneous factors should be for its own internally

developed model. In addition, increasing district participation and active involvement in model-building will also boost its feeling of "ownership" of the research, as opposed to the often-heard complaint of having such research "imposed" by outsiders.)

Finally, as a result of all of the above, "external validity" (or generalizability) of the model ought to be greatly enhanced. This is because care has been taken to incorporate those measures which are particularly relevant and unique to each district.

To summarize, the first goal of the analysis will be to assess the magnitude and direction of gain scores for selected student achievement measures within individual districts. Secondly, a predictive model will be developed within selected districts which links multiple district-specific student achievement and teacher performance measures, while controlling for possible pre-existing extraneous influences on student performance.

## References


Baker, T. L. (1988). Doing social research. New York: McGraw-Hill Company.

Dereshiwsky, M. I., Packard, R. D. (1988, April 21) Predictive achievement model. Paper presented to the Career Ladder Steering Committee, Sunnyside School District, Tucson, Arizona.

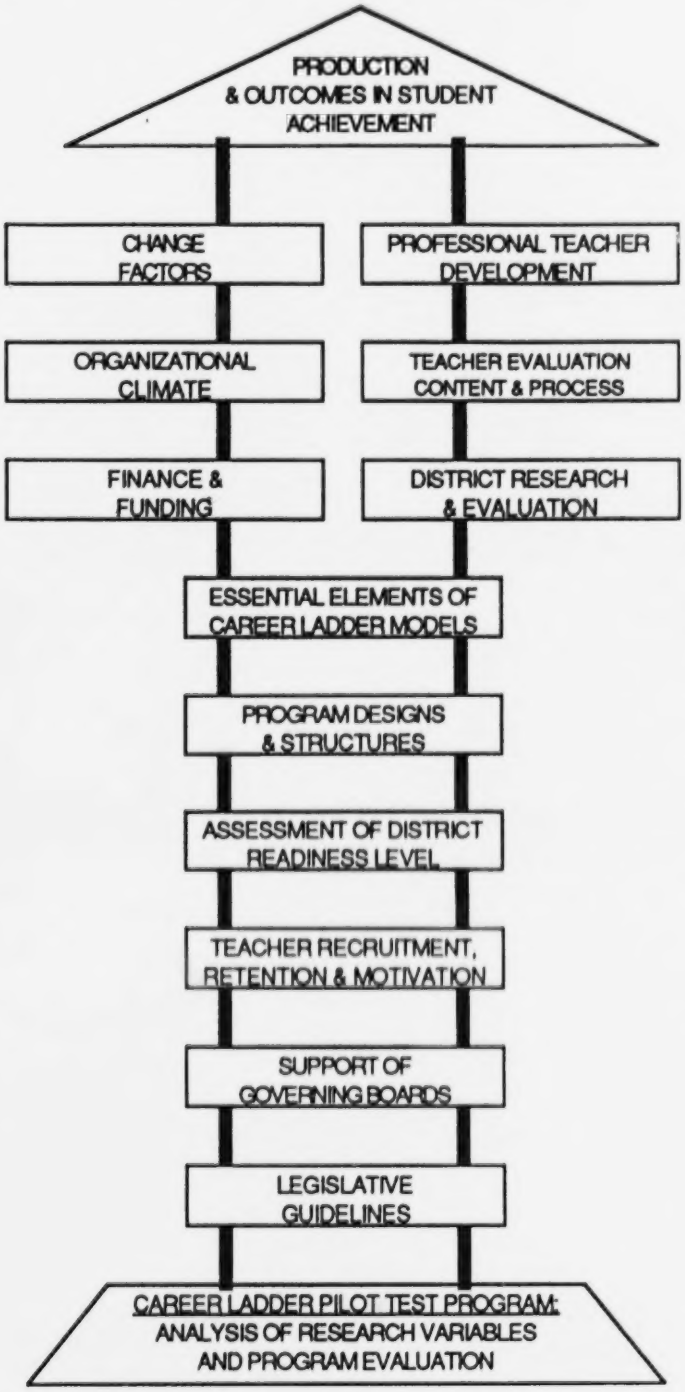
Helmstadter, G. C. (1987). An index of teacher effectiveness which is fair, objective and based on student performance. Paper presented to the Arizona Joint Legislative Committee on Career Ladders, Arizona State University, Tempe, Arizona.

Packard, R. D. (1988, March 24). University & pilot career ladder district collaborative planning. Document presented to the Career Ladder Pilot District Network, Mesa, Arizona.

Packard, R. D. (1988, April 18). Preparation of the network task force decision package on research content and procedures. Document presented to the Career Ladder Pilot District Task Force, Mesa, Arizona.

APPENDIX A

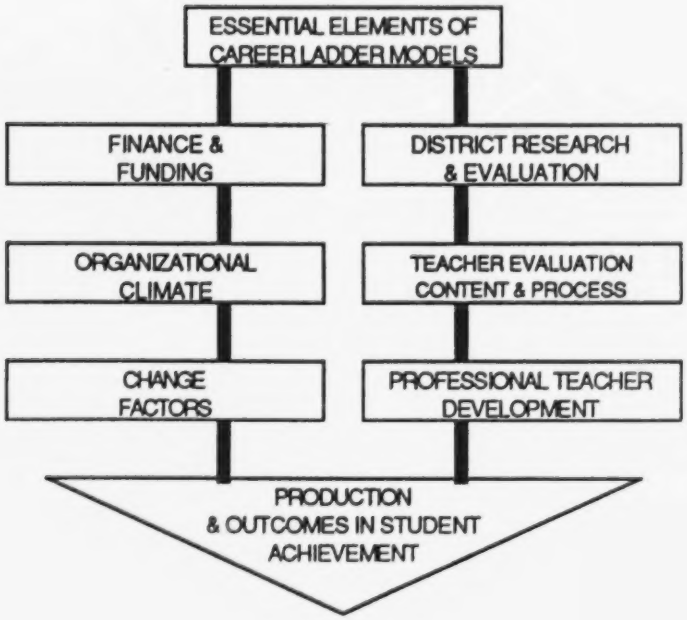
## EDUCATIONAL REFORM AND DEVELOPMENT IN THE STATE OF ARIZONA


PRODUCTION & OUTCOMES IN STUDENT ACHIEVEMENT
CHANGE FACTORS
PROFESSIONAL TEACHER DEVELOPMENT
ORGANIZATIONAL CLIMATE
TEACHER EVALUATION CONTENT & PROCESS
FINANCE & FUNDING
DISTRICT RESEARCH & EVALUATION
ESSENTIAL ELEMENTS OF CAREER LADDER MODELS
PROGRAM DESIGNS & STRUCTURES
ASSESSMENT OF DISTRICT READINESS LEVEL
TEACHER RECRUITMENT, RETENTION & MOTIVATION
SUPPORT OF GOVERNING BOARDS
LEGISLATIVE GUIDELINES
CAREER LADDER PILOT TEST PROGRAM:
ANALYSIS OF RESEARCH VARIABLES AND PROGRAM EVALUATION

CENTER FOR EXCELLENCE IN EDUCATION
NORTHERN ARIZONA UNIVERSITY

APPENDIX B

## TEACHER INCENTIVE PROGRAM COMPONENTS FOR ENHANCED STUDENT DEVELOPMENT




CENTER FOR EXCELLENCE IN EDUCATION
NORTHERN ARIZONA UNIVERSITY

# END

U.S. DEPT. OF EDUCATION

OFFICE OF EDUCATIONAL RESEARCH AND IMPROVEMENT (OERI)

# ERIC

DATE FILMED

# AUG 15 1988